

08/182,183
Part 25

(1) GENERAL INFORMATION

(i) APPLICANT: LIN, LEU-FEN
COLLINS, FRANKLIN D.
DOHERTY, DANIEL H.
LILE, JACK
BEKTESH, SUSAN

(ii) TITLE OF INVENTION: Glial Cell Line-Derived Neurotrophic Factor

(iii) NUMBER OF SEQUENCES: 25

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: AMGEN INC.
(B) STREET: 1840 DeHavilland Drive
(C) CITY: Thousand Oaks
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 91320-1789

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy Disk
(B) COMPUTER: Macintosh
(C) OPERATING SYSTEM: 7.1
(D) SOFTWARE: Microsoft Word 5.1

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/182,183
(B) FILING DATE: 23-MAY-1994

(2) INFORMATION FOR SEQ ID NO:1

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(C) TOPOLOGY: linear

(v) FRAGMENT TYPE: N-terminal fragment

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ser Pro Asp Lys Gln Ala Ala Ala Leu Pro Arg Arg Glu Arg Asn Xaa
5 10 15

Gln Ala Ala Ala Ala Ser Pro Asp Asn
20 25

(2) INFORMATION FOR SEQ ID NO:2

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) TOPOLOGY: linear

(v) FRAGMENT TYPE: internal fragment

(ix) FEATURE: Xaa is either Lys or Gln

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Asp Xaa Ile Leu Lys Asn Leu Gly Arg Val Arg Arg Leu
5 10

(2) INFORMATION FOR SEQ ID NO:3

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 900 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: nucleic acid for rat GDNF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

| | |
|---|-----|
| CCCCCGGGCT GCAGGAATTC GGGG GTC TAC GGA GAC CGG ATC CGA GGT Val Tyr Gly Asp Arg Ile Arg Gly -90 | 48 |
| GCC GCC GCC GGA CGG GAC TCT AAG ATG AAG TTA TGG GAT GTC GTG Ala Ala Ala Gly Arg Asp Ser Lys Met Lys Leu Trp Asp Val Val -85 -80 -75 | 93 |
| GCT GTC TGC CTG GTG TTG CTG CAC ACC GCG TCT GCC TTC CCG CTG Ala Val Cys Leu Val Leu Leu His Thr Ala Ser Ala Phe Pro Leu -70 -65 -60 | 138 |
| CCC GCC GGT AAG AGG CTT CTC GAA GCG CCC GCC GAA GAC CAC TCC Pro Ala Gly Lys Arg Leu Leu Glu Ala Pro Ala Glu Asp His Ser -55 -50 -45 | 183 |
| CTC GGC CAC CGC CGC GTG CCC TTC GCG CTG ACC AGT GAC TCC AAT Leu Gly His Arg Arg Val Pro Phe Ala Leu Thr Ser Asp Ser Asn -40 -35 -30 | 228 |
| ATG CCC GAA GAT TAT CCT GAC CAG TTT GAT GAC GTC ATG GAT TTT Met Pro Glu Asp Tyr Pro Asp Gln Phe Asp Asp Val Met Asp Phe -25 -20 -15 | 273 |
| ATT CAA GCC ACC ATC AAA AGA CTG AAA AGG TCA CCA GAT AAA CAA Ile Gln Ala Thr Ile Lys Arg Leu Lys Arg Ser Pro Asp Lys Gln -10 -5 1 5 | 318 |
| GCG GCG GCA CTT CCT CGA AGA GAG AGG AAC CGG CAA GCT GCA GCT Ala Ala Ala Leu Pro Arg Arg Glu Arg Asn Arg Gln Ala Ala Ala 10 15 20 | 363 |
| GCC AGC CCA GAG AAT TCC AGA GGG AAA GGT CGC AGA GGC CAG AGG Ala Ser Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg 25 30 35 | 408 |

| | |
|---|-----|
| GGC AAA AAT CGG GGG TGC GTC TTA ACT GCA ATA CAC TTA AAT GTC | 453 |
| Gly Lys Asn Arg Gly Cys Val Leu Thr Ala Ile His Leu Asn Val | |
| 40 | 45 |
| 50 | |
| ACT GAC TTG GGT TTG GGC TAC GAA ACC AAG GAG GAA CTG ATC TTT | 498 |
| Thr Asp Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu Leu Ile Phe | |
| 55 | 60 |
| 65 | |
| CGA TAT TGT AGC GGT TCC TGT GAA GCG GCC GAG ACA ATG TAC GAC | 543 |
| Arg Tyr Cys Ser Gly Ser Cys Glu Ala Ala Glu Thr Met Tyr Asp | |
| 70 | 75 |
| 80 | |
| AAA ATA CTA AAA AAT CTG TCT CGA AGT AGA AGG CTA ACA AGT GAC | 588 |
| Lys Ile Leu Lys Asn Leu Ser Arg Ser Arg Arg Leu Thr Ser Asp | |
| 85 | 90 |
| 95 | |
| AAG GTA GGC CAG GCA TGT TGC AGG CCG GTC GCC TTC GAC GAC GAC | 633 |
| Lys Val Gly Gln Ala Cys Cys Arg Pro Val Ala Phe Asp Asp Asp | |
| 100 | 105 |
| 110 | |
| CTG TCG TTT TTA GAC GAC AGC CTG GTT TAC CAT ATC CTA AGA AAG | 678 |
| Leu Ser Phe Leu Asp Asp Ser Leu Val Tyr His Ile Leu Arg Lys | |
| 115 | 120 |
| 125 | |
| CAT TCC GCT AAA CGG TGT GGA TGT ATC TGA CCCTGGCTCC | 718 |
| His Ser Ala Lys Arg Cys Gly Cys Ile | |
| 130 | |
| AGAGACTGCT GTGTATTGCA TTCCCTGCTAC ACTGCGAAGA AAGGGACCAA | 768 |
| GGTTCCCAGG AAATATTTGC CCAGAAAGGA AGATAAGGAC CAAGAAGGCA | 818 |
| GAGGCAGAGG CGGAAGAAGA AGAAGAAAAG AAGGACGAAG GCAGCCATCT | 868 |
| GTGGGAGCCT GTAGAAGGAG GCCCAGCTAC AG | 900 |

(2) INFORMATION FOR SEQ ID NO:4

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 134 amino acid residues
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: inferred amino acid sequence for mature rat GDNF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

| | |
|---|----|
| Ser Pro Asp Lys Gln Ala Ala Ala Leu Pro Arg Arg Glu Arg Asn Arg | |
| 1 | 5 |
| 10 | 15 |

| | |
|---|----|
| Gln Ala Ala Ala Ala Ser Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg | |
| 20 | 25 |
| 30 | |

Gly Gln Arg Gly Lys Asn Arg Gly Cys Val Leu Thr Ala Ile His Leu
35 40 45

Asn Val Thr Asp Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu Leu Ile
50 55 60

Phe Arg Tyr Cys Ser Gly Ser Cys Glu Ala Ala Glu Thr Met Tyr Asp
65 70 75 80

Lys Ile Leu Lys Asn Leu Ser Arg Ser Arg Arg Leu Thr Ser Asp Lys
85 90 95

Val Gly Gln Ala Cys Cys Arg Pro Val Ala Phe Asp Asp Asp Leu Ser
100 105 110

Phe Leu Asp Asp Ser Leu Val Tyr His Ile Leu Arg Lys His Ser Ala
115 120 125

Lys Arg Cys Gly Cys Ile
130

(2) INFORMATION FOR SEQ ID NO:5

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 562 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: nucleic acid sequence for human GDNF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATTT TCTCTTTCT TTTTGAACAG CA AAT ATG CCA GAG GAT TAT CCT 47
Ser Asn Met Pro Glu Asp Tyr Pro
-25 -20

GAT CAG TTC GAT GAT GTC ATG GAT TTT ATT CAA GCC ACC ATT 89
Asp Gln Phe Asp Asp Val Met Asp Phe Ile Gln Ala Thr Ile
-15 -10

AAA AGA CTG AAA AGG TCA CCA GAT AAA CAA ATG GCA GTG CTT 131
Lys Arg Leu Lys Arg Ser Pro Asp Lys Gln Met Ala Val Leu
-5 1 5

CCT AGA AGA GAG CGG AAT CGG CAG GCT GCA GCT GCC AAC CCA 173
Pro Arg Arg Glu Arg Asn Arg Gln Ala Ala Ala Asn Pro
10 15 20

GAG AAT TCC AGA GGA AAA GGT CGG AGA GGC CAG AGG GGC AAA 215
Glu Asn Ser Arg Gly Lys Gly Arg Arg Gly Gln Arg Gly Lys
25 30 35

| | |
|---|-----|
| AAC CGG GGT TGT GTC TTA ACT GCA ATA CAT TTA AAT GTC ACT Asn Arg Gly Cys Val Leu Thr Ala Ile His Leu Asn Val Thr 40 45 50 | 257 |
| GAC TTG GGT CTG GGC TAT GAA ACC AAG GAG GAA CTG ATT TTT Asp Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu Leu Ile Phe 55 60 65 | 299 |
| AGG TAC TGC AGC GGC TCT TGC GAT GCA GCT GAG ACA ACG TAC Arg Tyr Cys Ser Gly Ser Cys Asp Ala Ala Glu Thr Thr Tyr 70 75 | 341 |
| GAC AAA ATA TTG AAA AAC TTA TCC AGA AAT AGA AGG CTG GTG Asp Lys Ile Leu Lys Asn Leu Ser Arg Asn Arg Arg Leu Val 80 85 90 | 383 |
| ACT GAC AAA GTA GGG CAG GCA TGT TGC AGA CCC ATC GCC TTT Ser Asp Lys Val Gly Gln Ala Cys Cys Arg Pro Ile Ala Phe 95 100 105 | 425 |
| GAT GAT GAC CTG TCG TTT TTA GAT GAT AAC CTG GTT TAC CAT Asp Asp Asp Leu Ser Phe Leu Asp Asp Asn Leu Val Tyr His 110 115 120 | 467 |
| ATT CTA AGA AAG CAT TCC GCT AAA AGG TGT GGA TGT ATC TGA Ile Leu Arg Lys His Ser Ala Lys Arg Cys Gly Cys Ile 125 130 | 509 |
| CTCCGGCTCC AGAGACTGCT GTGTATTGCA TTCCTGCTAC AGTGCAAAGA | 559 |
| AAG | 562 |

(2) INFORMATION FOR SEQ ID NO:6

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 134 amino acid residues
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: inferred amino acid sequence for mature human GDNF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

| |
|--|
| Ser Pro Asp Lys Gln Met Ala Val Leu Pro Arg Arg Glu Arg Asn Arg 1 5 10 15 |
|--|

| |
|---|
| Gln Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg Arg 20 25 30 |
|---|

| |
|---|
| Gly Gln Arg Gly Lys Asn Arg Gly Cys Val Leu Thr Ala Ile His Leu 35 40 45 |
|---|

Asn Val Thr Asp Leu Gly Leu Gly Tyr Glu Thr Lys Glu Glu Leu Ile
50 55 60

Phe Arg Tyr Cys Ser Gly Ser Cys Asp Ala Ala Glu Thr Thr Tyr Asp
65 70 75 80

Lys Ile Leu Lys Asn Leu Ser Arg Asn Arg Arg Leu Val Ser Asp Lys
85 90 95

Val Gly Gln Ala Cys Cys Arg Pro Ile Ala Phe Asp Asp Asp Leu Ser
100 105 110

Phe Leu Asp Asp Asn Leu Val Tyr His Ile Leu Arg Lys His Ser Ala
115 120 125

Lys Arg Cys Gly Cys Ile
130

(2) INFORMATION FOR SEQ ID NO:7

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: oligonucleotide probe
- (D) OTHER INFORMATION: N at positions 3, 15, and 18 is inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCNGAYAARC ARG CNGCNGC

20

(2) INFORMATION FOR SEQ ID NO:8

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 223 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: nucleic acid sequence for human GDNF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TTCTCTCCCC CACCTCCCGC CTGCCCGCGC A GGT GCC GCC GGC
Gly Ala Ala Ala Gly
-5

46

CGG GAC TTT AAG ATG AAG TTA TGG GAT GTC GTG GCT GTC TGC
Arg Asp Phe Lys Met Lys Leu Trp Asp Val Val Ala Val Cys
1 5 10

88

CTG GTG CTG CTC CAC ACC GCG TCC GCC TTC CCG CTG CCC GCC
Leu Val Leu Leu His Thr Ala Ser Ala Phe Pro Leu Pro Ala
15 20

130

GGT AAG AGG CCT CCC GAG GCG CCC GCC GAA GAC CGC TCC CTC 172
Gly Lys Arg Pro Pro Glu Ala Pro Ala Glu Asp Arg Ser Leu
25 30 35

GGC CGC CGC CGC GCG CCC TTC GCG CTG AGC AGT GAC T 209
Gly Arg Arg Arg Ala Pro Phe Ala Leu Ser Ser Asp
40 45 50

GTAAGAACCG TTCC 223

(2) INFORMATION FOR SEQ ID NO:9

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: linker

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCCGAATTCTG GG 12

(2) INFORMATION FOR SEQ ID NO:10

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Pro Asp Lys Gln Ala Ala Ala
5

(2) INFORMATION FOR SEQ ID NO:11

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: nucleic acid sequence from pBluescript
SK-76.1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GAGAGGAACC GGCAAGCTGC WGMWGYMWGM CCW 33

(2) INFORMATION FOR SEQ ID NO:12

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acid residues
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Glu Arg Asn Arg Gln Ala Ala Ala Ala Ser Pro
5 10

(2) INFORMATION FOR SEQ ID NO:13

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: oligonucleotide DHD-26
- (D) OTHER INFORMATION: N at positions 9 and 12 are inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ARRTTYTTNA RNATYTTRTC

20

(2) INFORMATION FOR SEQ ID NO:14

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7 amino acid residues
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Asp Lys Ile Leu Lys Asn Leu
5

(2) INFORMATION FOR SEQ ID NO:15

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: oligonucleotide primer PD1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GACGGGACTC TAAGATG

17

(2) INFORMATION FOR SEQ ID NO:16

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE

- (A) NAME/KEY: oligonucleotide primer DHD23

(D) OTHER INFORMATION: N at positions 3, 6, and 18 is inosine

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCNGCNGCYT GYTTRTCNGG

20

(2) INFORMATION FOR SEQ ID NO:17

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: oligonucleotide primer LF2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGAGACAATG TACCGACA

17

(2) INFORMATION FOR SEQ ID NO:18

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: oligonucleotide primer PD2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTCTGGAGCC AGGGTCA

17

(2) INFORMATION FOR SEQ ID NO:19

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: oligonucleotide primer PD1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCCGAATTCTG ACGGGACTCT AAGATG

26

(2) INFORMATION FOR SEQ ID NO:20

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: oligonucleotide primer LFA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGGTGGCCAG AGGGAGTGGT CTTC

24

(2) INFORMATION FOR SEQ ID NO:21

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: oligonucleotide primer PD3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGCGGATCCA ATAAGGAGGA AAAAAAATGT CACCAAGATAA ACAAAT

46

(2) INFORMATION FOR SEQ ID NO:22

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: oligonucleotide primer PD4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CGCGGTACCC AGTCTCTGGA GCCGGA

26

(2) INFORMATION FOR SEQ ID NO:23

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: adapter fragment for plasmid pCJ1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GATCTAGAAT TGTCTATGTTT GACAGCTTAT CAT

33

(2) INFORMATION FOR SEQ ID NO:24

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: polylinker sequence for plasmid pCJX1-1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24

AATTCCCGGG TACCAAGATCT GAGCTCACTA GTCTGCA

37

(2) INFORMATION FOR SEQ ID NO:25

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 747 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: nucleic acid sequence for human GDNF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25

| | |
|---|-----|
| TTCTCTCCCC CACCTCCCGC CTGCCCGCGC A GGT GCC GCC GGA | 46 |
| Gly Ala Ala Ala Gly | |
| CGG GAC TTT AAG ATG AAG TTA TGG GAT GTC GTG GCT GTC TGC | 88 |
| Arg Asp Phe Lys Met Lys Leu Trp Asp Val Val Ala Val Cys | |
| CTG GTG CTG CTC CAC ACC GCG TCC GCC TTC CCG CTG CCC GCC | 130 |
| Leu Val Leu Leu His Thr Ala Ser Ala Phe Pro Leu Pro Ala | |
| GGT AAG AGG CCT CCC GAG GCG CCC GCC GAA GAC CGC TCC CTC | 172 |
| Gly Lys Arg Pro Pro Glu Ala Pro Ala Glu Asp Arg Ser Leu | |
| GGC CGC CGC CGC GCG CCC TTC GCG CTG AGC AGT GAC TCA AAT | 214 |
| Gly Arg Arg Arg Ala Pro Phe Ala Leu Ser Ser Asp Ser Asn | |
| ATG CCA GAG GAT TAT CCT GAT CAG TTC GAT GAT GTC ATG GAT | 256 |
| Met Pro Glu Asp Tyr Pro Asp Gln Phe Asp Asp Val Met Asp | |
| TTT ATT CAA GCC ACC ATT AAA AGA CTG AAA AGG TCA CCA GAT | 298 |
| Phe Ile Gln Ala Thr Ile Lys Arg Leu Lys Arg Ser Pro Asp | |
| 1 | |
| AAA CAA ATG GCA GTG CTT CCT AGA AGA GAG CGG AAT CGG CAG | 340 |
| Lys Gln Met Ala Val Leu Pro Arg Arg Glu Arg Asn Arg Gln | |
| 5 10 15 | |
| GCT GCA GCT GCC AAC CCA GAG AAT TCC AGA GGA AAA GGT CGG | 382 |
| Ala Ala Ala Ala Asn Pro Glu Asn Ser Arg Gly Lys Gly Arg | |
| 20 25 30 | |
| AGA GGC CAG AGG GGC AAA AAC CGG GGT TGT GTC TTA ACT GCA | 424 |
| Arg Gly Gln Arg Gly Lys Asn Arg Gly Cys Val Leu Thr Ala | |
| 35 40 45 | |
| ATA CAT TTA AAT GTC ACT GAC TTG GGT CTG GGC TAT GAA ACC | 466 |
| Ile His Leu Asn Val Thr Asp Leu Gly Leu Gly Tyr Glu Thr | |
| 50 55 | |

| | |
|---|-----|
| AAG GAG GAA CTG ATT TTT AGG TAC TGC AGC GGC TCT TGC GAT Lys Glu Glu Leu Ile Phe Arg Tyr Cys Ser Gly Ser Cys Asp 60 65 70 | 508 |
| GCA GCT GAG ACA ACG TAC GAC AAA ATA TTG AAA AAC TTA TCC Ala Ala Glu Thr Thr Tyr Asp Lys Ile Leu Lys Asn Leu Ser 75 80 85 | 550 |
| AGA AAT AGA AGG CTG GTG ACT GAC AAA GTA GGG CAG GCA TGT Arg Asn Arg Arg Leu Val Ser Asp Lys Val Gly Gln Ala Cys 90 95 100 | 592 |
| TGC AGA CCC ATC GCC TTT GAT GAT GAC CTG TCG TTT TTA GAT Cys Arg Pro Ile Ala Phe Asp Asp Asp Leu Ser Phe Leu Asp 105 110 115 | 634 |
| GAT AAC CTG GTT TAC CAT ATT CTA AGA AAG CAT TCC GCT AAA Asp Asn Leu Val Tyr His Ile Leu Arg Lys His Ser Ala Lys 120 125 | 676 |
| AGG TGT GGA TGT ATC TGA CTCCGGCTCC AGAGACTGCT GTGTATTGCA Arg Cys Gly Cys Ile 130 | 724 |
| TTCCCTGCTAC AGTGCAAAGA AAG | 747 |